successfully but only after introducing a multiplicative scale parameter to distinguish between both approaches. With 1600 respondents, adding 1 TTO offers more informative value than adding 1 DC but not as much as adding 2 DC's. **CONCLUSIONS:** The likelihood approach effectively estimates the structure underlying the simulated data. Given that DC is less burdensome than TTO, one may prefer to add more DC's than TTO's. That is – as in this case – when the underlying modelling assumption apply.

UT2

UPDATE OF THE PATIENT-REPORTED OUTCOME AND QUALITY OF LIFE INSTRUMENTS DATABASE (PROQOLID) USING THE FDA GUIDANCE ON PRO MEASURES

Caron M, Perrier LL, Vaissier V, Savre I, Acquadro C, Emory MP
MAPI Research Trust, Lyon, France

OBJECTIVES: In 2002, PROQOLID was launched to provide an overview of existing PRO instruments and facilitate access to the instruments and their developers, through the structured presentation of synthesized, reliable, and updated data. In 2009, the Food and Drug administration (FDA) published its final guidance on the use of PRO measures which describes how the FDA will review PRO instruments used to support claims in approved medical product labeling. The objectives of this study were to review and adapt the template of PROQOLID to harmonize its structure and language with those used in the FDA guidance, acknowledging that PROQOLID and the guidance have different **OBJECTIVES:** provision of information vs. review of information. **METHODS:** Content and structure of PROQOLID and the FDA guidance were compared. Proposed changes in terminology and structure were submitted to a panel of PRO experts (n=2). **RESULTS:** The information on PROQOLID is divided into 12 sections. The FDA guidance categorizes information into 5 parts. Twelve changes in terminology were made across all sections. For instance, Time recall" was changed to "Recall/Observation period", and "Dimensions" to "Domains". Fourteen changes of structure were made, mainly in Sections 6 and 7. Section 6 (Methodology of development) was renamed "Content Validity documentation". In this section, the heading "Information retrieval" was replaced by "Concept elicitation and Item generation". "Conceptual framework" will be moved to Section 5 (Descriptive information). Section 7 (Psychometric properties) was renamed "Measurement Properties". Within the "Reliability" heading, a subheading on "Inter-interviewer reproducibility" was added. A new section was created: "Data analysis and Interpretation". Five sections remained unchanged (1 to 4, and 8). **CONCLUSIONS:** The comparison of PROQOLID and the FDA guidance led to numerous changes in the wording and structure of the database. These changes will improve the functionality of PROQOLID and help users to better fulfill FDA requirements.

UT3

THE VALIDITY OF THE EQ-5D, SF-6D, SF-36 AND SF-12 IN MENTAL HEALTH CONDITIONS: A SYSTEMATIC REVIEW

Papaioannou D, Peasgood T, Brazier J, Parry G
University of Sheffield, Sheffield, UK

OBJECTIVES: To assess the validity and responsiveness of the SF-36, SF-12, SF-6D and EQ-5D in mental health conditions. **METHODS:** Systematic reviews were undertaken in five mental health conditions. Ten databases were searched to August 2009. Studies were appraised and data extracted. A narrative synthesis was performed on construct validity including known groups validity (KGV) (ability to detect differences in HRQL scores between groups), convergent validity (CV) (strength of association between generic HRQL and other related measures (e.g. symptoms or function) and responsiveness (R) (i.e. changes in scores in responders/non-responders to treatment and correlation with changes in related measures). **RESULTS:** Within schizophrenia, the majority of evidence related to the SF-36 (n=25) and EQ-5D (n=9). Both measures demonstrated KGV but this was mostly limited to demonstrating differences between individuals with schizophrenia and the general population. Contradictory results were found in studies measuring CV and R using clinical measures of symptom severity. For bipolar disorder, 23 studies were identified, almost exclusively on the SF-36; which was able to detect known differences in symptom severity and correlated strongly with clinical measures of depression (weakly with mania measures). For personality disorders; the majority of studies (6/9) related to the EQ-5D, which showed good KGV and R. For depression and anxiety, 23 EQ-5D and eight SF-6D studies were identified. Both measures demonstrated good CV and R for depression; however KGV may be driven by the presence of co-morbid depression in patients with anxiety disorders. **CONCLUSIONS:** Overall, the evidence suggests that the generic HRQL measures are appropriate in four mental health conditions, but raises doubts about their use in schizophrenia. Caution is required when interpreting CV evidence using clinical measures, since the lack of relationship may reflect genuine lack of difference in HRQL. More evidence using better indicators for testing validity and responsiveness are required.

UT4

COMPARISON OF THE PERFORMANCE OF EQ-5D AND SF-6D IN PATIENTS WITH CHRONIC PAIN –RESULTS FROM 3 RANDOMIZED CONTROLLED TRIALS

Obradovic M, Lal A, Liedgens H
Grünenthal GmbH, Aachen, Germany

OBJECTIVES: To compare EQ-5D and SF-6D utilities in patients with chronic pain due to osteoarthritis (OA) of the knee or low back pain (LBP) derived from phase III trials with tapentadol. **METHODS:** Three phase III trials with identical design included EQ-5D and SF-36 questionnaires to measure quality of life of patients with pain due to OA or LBP treated with tapentadol prolonged release (PR), oxycodone controlled release (CR) or placebo. EQ-5D and SF-6D indices were obtained using the UK weights. The ability of the two utility measurements to discriminate among different health states was tested. **RESULTS:** Both SF-6D and EQ-5D utility values increased from baseline to endpoint (15 weeks). The increase (mean of all patients with active treatment) was substantially higher when measured with EQ-5D (0.16 vs. 0.06). The EQ-5D better distinguished among health states (different severity of adverse events, pain relief, withdrawal rates). While utilities were very similar in a group of patients who tolerated the treatment (0.695 and 0.694 for EQ-5D and SF-6D, respectively), EQ-5D utilities were considerably lower in patients who withdrew due to adverse events (0.503 and 0.597 for EQ-5D and SF-6D, respectively). A similar pattern was seen in patients with various levels of pain relief. In patients with >30% pain relief mean EQ-5D and SF-6D utility was 0.716 and 0.708, respectively. The EQ-5D utility in patients who withdrew due to lack of efficacy was 0.405, when analyzing the SF-6D utility this resulted in 0.580. **CONCLUSIONS:** Both generic instruments to measure quality of life, EQ-5D and SF-6D, showed that avoidance of severe treatment-related adverse events and sufficient pain relief has a large beneficial impact on patient's wellbeing. In the clinical trials analyzed the discriminative power of the EQ-5D was stronger showing that this instrument is a useful tool also in pain studies to analyze patient's QoL.

PODIUM SESSION II: DISCUSSIONS ON THE ADDED VALUE OF VALUE OF INFORMATION

VI1

DETERMINING THE IMPACT OF MODELING ADDITIONAL SOURCES OF UNCERTAINTY IN VALUE OF INFORMATION ANALYSIS

Corro Ramos I, Rutten-van mölken MP, Al MJ
Erasmus University, Rotterdam, The Netherlands

The conditional reimbursement policy for expensive medicines in the The Netherlands requires real-world data collection on cost-effectiveness within a four years period (T=4) after the initial decision to reimburse a drug (T=0). This introduces *new sources of uncertainty*, which are less important in an RCT than in real life. This may affect the priorities for further (real-world) research as determined in a VOI analysis. **OBJECTIVES:** Identifying and modeling types of uncertainty that are usually not parameterized at T=0 but may become relevant at T=4. Include them in the VOI analysis. **METHODS:** We use a hypothetical model with four states and parameters related to transition and exacerbation probabilities, costs and utilities. Three additional uncertainties were parameterized: persistence, compliance and broadening of indication. Persistence refers to the duration of the treatment and it is determined by the probability of dropping out of the treatment. Compliance is characterized by the fraction of the treatment benefit obtained due to not taking the medication as it was indicated. The impact of indication broadening is modeled as the percentage of the RCT treatment effect realized in the outcome study. These extra parameters were included in the VOI analysis. **RESULTS:** Priorities change when new uncertainties are introduced in the model. Initially, the EVPPI was highest for transition probabilities followed by utilities; and it was very low for exacerbation probabilities and costs. After new uncertainties are included, compliance and broadening of indication (which is applied only to the new treatment) become as relevant as utilities. Persistence however has little impact in the model. **CONCLUSIONS:** VOI analysis at T=0 should anticipate and parameterize new types of uncertainty that may emerge during a four year outcomes study. This would help to focus the real-world outcomes study on those parameters that reduce uncertainty in the decision to continue the reimbursement most.

VI2

A NOVEL APPROACH TO ANALYSING VALUE DRIVER IMPORTANCE ACROSS MULTIPLE TARGET PRODUCT PROFILES

Wild L[1], Mukku SR[2]
[1]*Double Helix Consulting, London, UK,* [2]*Double Helix Consulting Group, London, UK*

OBJECTIVES: To develop a novel semi-quantitative model to assist pharmaceutical companies in making investment decisions, by assessing relative importance of value drivers in a given therapy area and how this translates to perceived value of a product profile. **METHODS:** Perceived value for a number of product profiles is assessed through a semi-quantitative scoring method, followed by an in-depth qualitative interview. In the scoring phase, respondents rate the relative importance of value drivers and provide thresholds for minimal and strong value in each domain. Respondents assess target product profiles, scoring profile performance in each value driver on a pre-defined scale. **RESULTS:** This methodology provides a range of valuable data in understanding the drivers of value in a given therapy area. First, the relative importance of value drivers can be used to understand which product attributes (efficacy, safety and tolerability or administration and others) drive product value. In addition, by providing value thresholds for each driver, we can understand expectations, in effect defining an 'ideal' product scenario. Testing product profiles against a scale calibrated by these expectations allows us to understand perceived product value in a set of likely product attributes. In addition, by testing a number of profiles, trade-offs between different product attributes, and the effect of these on product value, can be assessed. **CONCLUSIONS:** The insights gained from this type of analysis are vital in understanding product development priorities and the likely pricing and reimbursement potential for future products. Multiple applications of this technique have confirmed that this is a valuable approach in supporting pharmaceutical companies to inform their clinical programme, pricing and reimbursement strategy or commercial strategy.

VI3

SEQUENTIAL TREATMENT OF FOLLICULAR NON-HODGKIN LYMPHOMA: COST-EFFECTIVENESS AND VALUE OF INFORMATION

Soini EJ[1], Martikainen JA[1], Vihervaara V[2], Mustonen K[2], Nousiainen T[3]
[1]*ESiOR Ltd, Kuopio, Finland,* [2]*Roche Oy, Espoo, Finland,* [3]*Department of Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland*

OBJECTIVES: To assess the lifetime incremental cost-effectiveness ratios (ICER) per quality-adjusted life-year (QALY) gained and multinomial expected value of perfect information (mEVPI) of sequential follicular lymphoma (FL) treatment in Finland. **METHODS:** The novel cancer treatments included rituximab (R) and bendamustine (B). A probabilistic Markov-model was developed to simulate patients' transitions between first-line progression-free (PF1), PF2, progression and death states using a second-order Monte Carlo-simulation, one month cycle, and half cycle correction. All patients received the recommended induction with R-cyclophosphamide-doxorubicin-vincristine-prednisone (RCHOP). For the RCHOP-induction responders, the sequence was continued without the first-line R-maintenance treatment (RCHOP) or with it (RCHOPR). PF1 was based on the best fitting parametric extrapolation (Gompertz; 4-year treatment benefit trunk) of PRIMA (Primary RItuximab and MAintenance) data. After RCHOPR or RCHOP, eligible patients were assigned to second-line RCOPR/B or RCOPR/COP based on the PRIMA results, B indication/labelling and the recent ESMO (European Society for Medical Oncology) guideline for FL. PF2s (5-year treatment benefit trunk) were based on the parametric estimate of EORTC20981 and adjustment based on Rummel's trial. After PF2 (progression), patients received best supportive care (BSC). Age-dependent death was set equal to the larger of EORTC20981 or Finnish background mortality. Payer costs were included in 2010 value, and the most affordable public drug costs (2/2011; wastage included) were used. EQ-5D-based utilities were set 0.78 for PF1/PF2 and 0.62 for progression. 3% annual discounting was used. **RESULTS:** The ICERs for RCHOPR->RCOPR/B->BSC, RCHOPR->RCOPR/COP->BSC and RCHOP->RCOPR/B->BSC were €9575, €9881 and €8812 per QALY gained in comparison to RCHOP->RCOPR/COP->BSC, respectively. According to the cost-effectiveness acceptability frontier, 47% of patients with RCHOP->RCOPR/COP->BSC, 46% and 68% of patients with RCHOPR->RCOPR/B->BSC were cost effective at the ICER-levels of €5,000 (mEVPI €5,047/patient), €15,000 (mEVPI €3,101/patient) and €25,000 (mEVPI €1,564/patient) per QALY gained, respectively. **CONCLUSIONS:** First-line R-maintenance is an efficient and potentially cost-effective start for FL-treatment sequence.

VI4

THE VALUE OF INFORMATION OF A MULTICENTRE RANDOMISED CONTROLLED TRIAL OF INTRAVENOUS IMMUNOGLOBULIN FOR SEPSIS (SEVERE SEPSIS AND SEPTIC SHOCK)

Soares MO
University of York, York, UK

OBJECTIVES: Sepsis is a syndrome characterised by a systemic inflammatory response to infection. Intravenous immunoglobulin (IVIG) has been proposed as an adjuvant therapy for sepsis. The need for a high quality randomised controlled trial (RCT) to evaluate the effectiveness of this treatment has been identified in the literature. The objectives of the current work were to use value of information analyses (VOI) to establish whether the costs of carrying out an RCT are outweighed by the potential benefit of the resulting information. **METHODS:** A decision model was developed to evaluate the cost-effectiveness of IVIG (in terms of cost per quality-adjusted life-year) in adults with severe sepsis. A systematic review coupled with formal modelling selection process was performed to assess evidence on the relative effectiveness of IVIG. Further reviews were conducted to inform other relevant model parameters. Decision uncertainty was presented and the value of information assessed. **RESULTS:** A large degree of between study heterogeneity in the existing evidence base over the relative effectiveness of IVIG could be explained by a measure of study quality and duration of IVIG therapy. When using this model, the incremental cost-effectiveness ratio of IVIG was estimated to be £20,850 per QALY (threshold of £20,000 per QALY), and the probability of IVIG being cost effective was 0.505. No clear clinical rationale for the association between relative effectiveness and duration of therapy emerged from existing studies. Alternative models were evaluated for their impact on cost effectiveness and on the need for further research. The results demonstrate that conclusions are highly sensitive to the choice of model used for clinical effectiveness. **CONCLUSIONS:** Although the analyses suggested potential value from a large multicentre RCT, the uncertainties around the design of such a study mean that further dose-ranging/finding studies should conducted prior to funding any future multicentre RCT.

PODIUM SESSION III:
TAKING HETEROGENEITY INTO ACCOUNT BETWEEN PATIENTS AND BETWEEN STUDIES

HG1

APPLYING FRAILTY MODEL IN LONGITUDINAL SURVIVALS OF CHRONIC DISEASES

Li Y, Holtzer-Goor K, Uyl-de Groot C, Al M
iMTA, Erasmus University Rotterdam, Rotterdam, The Netherlands

OBJECTIVES: Survival analysis plays an important role in assessing the effectiveness of medical product/treatments and the risk factors. Particularly, the accelerated failure time (AFT) model provides the hazard/survival functions to the later health economic decision model to compute the cost-effectiveness. The model estimations of the usual survival models rely on the maximum likelihood estimation (MLE), which works under the assumption of independence of observations. However, this assumption is not always satisfied, especially in the chronic/relapsing disease. A patient with a chronic disease may experience recurrent disease progressions in the life course. When the progression free survivals (PFS) are recorded longitudinally, the PFS of different patients can be considered independent. Nevertheless, owing to sharing some unobserved heterogeneity, PFS of the same patients tend to associate with each other. This within-patient association can affect the estimation accuracy, therefore may misinform the decision makers. Particularly designed for the multivariate survival analysis, the frailty model takes this issue into account. **METHODS:** Fist, in a simulation study, we compared the AFT model and the Frailty model, where 5000 hypothetical patients are assigned to two treatment arms, and each patient experiences 5 treatment lines. Second, we apply both the Weibull AFT model and the Weibull-Gamma frailty model to the real life data, where 254 patients of chronic lymphocytic leukemia(CLL) have been followed, and we conduct a hypothesis test on the significance of frailty term. **RESULTS:** The simulation study shows that the estimates of the AFT model deviate far from the true values when the unobserved heterogeneity is large. The real life study indicates that the AFT model should be replaced by the frailty model due to the significance of the frailty term. **CONCLUSIONS:** In modelling survivals for chronic disease, the frailty models provide more accurate effect estimation than the conventional survival model.

HG2

CHARACTERIZING THE INDIVIDUAL COURSE OF HEALTH-RELATED QUALITY OF LIFE AFTER SUBARACHNOID HEMORRHAGE

Winter Y[1], Klotsche J[2], Ringel F[3], Spottke A[4], Gharevi N[4], Klockgether T[4], Schramm J[4], Urbach H[4], Wittchen HU[2], Rehm J[2], Meyer B[3], Dodel R[1]
[1]*Philipps-University, Marburg, Germany,* [2]*Technical University Dresden, Dresden, Germany,* [3]*Klinikum rechts der Isar, Technical University, Munich, Germany,* [4]*University of Bonn, Bonn, Germany*

OBJECTIVES: Subarachnoid hemorrhage (SAH) is a cerebrovascular disease leading to severe disability. SAH-patients show heterogeneous profiles of health-related quality of life (HrQoL). No methodological approaches to characterize the individual course of HrQoL following SAH have been developed. The objective was to characterize individual trajectories concerning HrQoL following SAH by using latent growth mixture modeling (LGMM). **METHODS:** A total of 113 incident patients with aneurysmal SAH treated in the University of Bonn between January 2004 and December 2005 were recruited in this longitudinal study. Clinical parameters (Hunt and Hess scale, Barthel-Index, Rankin Scale, Beck Depression Inventory) and HrQoL data (EQ-5D) were evaluated at baseline, 6 and 12 months. LGMM was applied to analyse the heterogeneity in individual courses of HrQoL after SAH. **RESULTS:** We identified four subgroups of patients (latent classes) with different patterns of HrQoL-course. Class 1 had the worst HrQoL-course with a low score of the EQ-5D index at baseline (0.33) and a non-significant change in scores over time. Patients in class 3 showed rapid recovery from initially low EQ-5D scores (0.37) during the first 6 months (D=0.47). Patients in classes 2 and 4 had 48-57% better initial HrQoL and similarly high HrQoL scores after 12 months. Patients in class 4 experienced a temporary reduction of HrQoL by 55%. The following clinical parameters were identified to characterize differences between classes: severity of SAH (Hunt and Hess scale), functional outcome, cognitive impairment and post-stroke depression. Treatment of post-stroke depression in classes 1 and 4 can improve HrQoL measures by factor 1.3-2.8. **CONCLUSIONS:** This methodological approach can be applied for more elaborated understanding of individual differences in long-term course of HrQoL after SAH. Identification of different patterns of disease course using LGMM may help to find subgroups of treatment responders and to assist the development of individual therapy regimes.

HG3

COST COMPARISONS AND METHODOLOGICAL HETEROGENEITY IN COST-OF-ILLNESS STUDIES: THE EXAMPLE OF COLORECTAL CANCER

Hanly P[1], O'Ceilleachair A[2], Skally M[3], O'Neill C[4], Sharp L[3]
[1]*National Cancer Registry, Cork, Co. Cork, Ireland,* [2]*National Cancer Registry, Cork, Cork, Ireland,* [3]*National Cancer Registry Ireland, Cork, Cork, Ireland,* [4]*National University of Ireland Galway, Galway, Galway, Ireland*

OBJECTIVES: Colorectal cancer (CRC) is the third most commonly-diagnosed cancer worldwide with over one million new cases globally each year. Advances in treatment and survival have been occurring, which is likely to have increased lifetime costs of managing the disease. We systematically reviewed and critiqued the cost-of-illness (COI) literature on CRC. **METHODS:** We searched Medline, EMBASE, CRD and the Cochrane library for CRC COI studies published in English, January 2000- February 2011. Data was abstracted independently by two reviewers on: setting, patient population, top-down/bottom-up costs, incident/prevalent approach, payer perspective, time horizon, costs included, cost source and per-person costs. We developed a framework to compare study methodologies and assess homogeneity/heterogeneity. **RESULTS:** Twenty-six papers met the inclusion criteria from the US (17), France (3), the UK (2), Canada (2), Switzerland (1) and Taiwan (1). Extensive methodological heterogeneity existed between studies. 17 studies employed top-down costings; 6 studies were prevalent, 8 incident and the remainder mixed. Time horizons ranged from 1-year post-diagnosis to lifetime. 25 studies included healthcare payer direct medical costs; 2 included indirect costs; 1 considered patient costs. The included papers described case-control studies based on claims/reimbursement data(10), examinations of patient charts (5) and analysis of claims data(4). There was broad agreement in how studies accounted for time, but few studies described costs in sufficient detail to allow repeatability. In general costs were not comparable between studies. There were some commonality in findings between studies from the same setting and which estimated costs in the year following diagnosis, but estimates varied greatly for longer time horizons. **CONCLUSIONS:** Methodological heterogeneity and lack of transparency made it almost impossible to compare CRC costs between studies or over time. For COI studies to be more informative, and amenable to external comparison, researchers, decision-makers and funders should adopt more standardised methodologies and promote greater transparency.